

99 203

YOR

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 9月10日

出 願 番 号

Application Number:

平成11年特許願第257587号

出 願 人

Applicant (s):

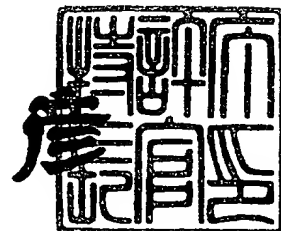
インターナショナル・ビジネス・マシーンス・コーポレイション

CERTIFIED COPY OF
PRIORITY DOCUMENT

1999年11月12日

特許庁長官
Commissioner,
Patent Office

近 藤 隆 彦



【書類名】 特許願

【整理番号】 JA999203

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 13/00

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 大和事業所内

【氏名】 田原 義則

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 大和事業所内

【氏名】 友田 大輔

【特許出願人】

【識別番号】 390009531

【住所又は居所】 アメリカ合衆国 1 0 5 0 4、ニューヨーク州アーモンク
(番地なし)

【氏名又は名称】 インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

【識別番号】 100086243

【弁理士】

【氏名又は名称】 坂口 博

【連絡先】 0 4 6 2 - 7 3 - 3 3 1 8、3 3 2 5、3 4 5 5

【選任した代理人】

【識別番号】 100091568

【弁理士】

【氏名又は名称】 市位 嘉宏

【手数料の表示】

【予納台帳番号】 024154

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9304391

【包括委任状番号】 9304392

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声情報の登録方法、認識文字列の特定方法、音声認識装置、音声情報の登録のためのソフトウェア・プロダクトを格納した記憶媒体、及び認識文字列の特定のためのソフトウェア・プロダクトを格納した記憶媒体

【特許請求の範囲】

【請求項 1】

音声入力装置を備える音声認識装置上で実施される音声情報の登録方法であって、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する段階と、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する段階と、

(c) 前記単語の読みに基づいてベースフォームを取得する段階と、

(d) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録する段階と、

を含む音声情報の登録方法。

【請求項 2】

音声入力装置を備える音声認識装置上で実施される認識文字列の特定方法であって、

(a 1) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する段階と、

(a 2) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する段階と、

(a 3) 前記単語の読みに基づいてベースフォームを取得する段階と、

(a 4) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録する段階と、

を含む登録段階と、

(b 1) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する段階と、

(b2) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する段階と、

(b3) 前記音声認識文字列を前記第 i の認識対象文字列と比較する段階と、
を含む認識段階と、
を含む認識文字列の特定方法。

【請求項 3】

前記認識対象文字列群は、アプリケーションから取得されたものであり、前記第 i の認識対象文字列に対応した制御メッセージを生成し、前記アプリケーションに送信する段階をさらに含む請求項 2 に記載の認識文字列の特定方法。

【請求項 4】

前記単語の読みには読みスコアが関連付けられて保存されており、また、前記ベースフォームには発音スコアが関連付けられて保存されており、前記読みスコアと前記発音スコアから導かれる関数値が閾値を越えている場合に前記ベースフォームが前記音声辞書に登録されることを特徴とする請求項 2 に記載の認識文字列の特定方法。

【請求項 5】

音声入力装置を備える音声認識装置上で実施される認識文字列の特定方法であって、

(a1) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する段階と、

(a2) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する段階と、

(a3) 前記単語の読みに基づいてベースフォームを取得する段階と、

(a4) 前記ベースフォームのスコアを計算する段階と、

(a5) 前記ベースフォームのスコアが閾値を越えている場合、前記単語に対応付けて音声認識辞書に登録する段階と、

を含む登録段階と、

(b1) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する段階と、

(b 2) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する段階と、

(b 3) 前記音声認識文字列を前記第 i の認識対象文字列と比較する段階と、

(b 4) 前記音声認識文字列と前記第 i の認識対象文字列が完全に一致した場合には、完全に一致した文字列に関連する処理を実行する段階と、

(b 5) 前記音声認識文字列の一部と前記第 i の認識対象文字列の一部が一致した場合には、前記閾値を低い値に更新し、第 2 の閾値とする段階と、

(b 6) 前記第 i の認識対象文字列に含まれる未知語を検出する段階と、

(b 7) 検出された未知語の読みを取得する段階と、

(b 8) 前記未知語の読みに基づいて第 2 のベースフォームを取得する段階と、

(b 9) 前記第 2 のベースフォームのスコアを計算する段階と、

(b 10) 前記第 2 のベースフォームのスコアが第 2 の閾値を越えている場合、前記未知語に対応付けて前記音声認識辞書に登録する段階と、
を含む認識段階と、

を含む認識文字列の特定方法。

【請求項 6】

音声入力装置を備える音声認識装置であって、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する認識文字列特定部と、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する未知語検出部と、

(c) 前記単語の読みに基づいてベースフォームを取得するベースフォーム生成部と、

(d) 前記ベースフォームが前記単語に対応付けて格納される音声認識辞書と

を含む音声認識装置。

【請求項 7】

音声入力装置を備える音声認識装置であって、

(a 1) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する認識文字列特定部と、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する未知語検出部と、

(c) 前記単語の読みに基づいてベースフォームを取得するベースフォーム生成部と、

(d) 前記ベースフォームを前記単語に対応付けて格納する音声認識辞書と、

(e) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する音声入力部と、

(f) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する音声認識エンジンと、

を含み、

(a 2) 前記認識特定部は、前記音声認識文字列を前記第 i の認識対象文字列と比較する音声認識装置。

【請求項 8】

前記認識文字列特定部は、前記認識対象文字列群を、アプリケーションから取得し、前記第 i の認識対象文字列に対応した制御メッセージを生成し、前記アプリケーションに送信することを特徴とする請求項 7 に記載の音声認識装置。

【請求項 9】

前記単語の読みには読みスコアが関連付けられて保存されており、また、前記ベースフォームには発音スコアが関連付けられて保存されており、前記読みスコアと前記発音スコアから導かれる関数値が閾値を越えている場合に前記ベースフォームが前記音声辞書に登録されることを特徴とする請求項 7 に記載の音声認識装置。

【請求項 10】

音声入力装置を備える音声認識装置であって、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する認識文字列特定部と、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認

識辞書には存在しない単語の読みを取得する未知語検出部と、

(c) 前記単語の読みに基づいてベースフォームを取得し、前記ベースフォームのスコアを計算するベースフォーム生成部と、

(d) 前記ベースフォームのスコアが閾値を越えている場合、前記単語に対応付けて前記ベースフォームを登録する音声認識辞書と、

(e) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する音声入力部と、

(f) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する音声認識エンジンと、

を含み、

(a 2) 前記認識文字列特定部は、

前記音声認識文字列を前記第 i の認識対象文字列と比較し、前記音声認識文字列と前記第 i の認識対象文字列が完全に一致した場合には、完全に一致した文字列に関連する処理を実行し、前記音声認識文字列の一部と前記第 i の認識対象文字列の一部が一致した場合には、前記閾値を低い値に更新し、第 2 の閾値とし、前記未知語検出部に第 i の認識対象文字列に含まれる未知語を検出し、検出された未知語の読みを取得することを指示し、前記未知語の読みに基づいて第 2 のベースフォームを取得し、前記第 2 のベースフォームのスコアを計算することを前記ベースフォーム生成部に指示し、前記第 2 のベースフォームのスコアが第 2 の閾値を越えている場合、前記未知語に対応付けて前記音声認識辞書に登録する音声認識装置。

【請求項 11】

音声入力装置を備える音声認識装置上で実施される音声情報の登録処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認

識辞書には存在しない単語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記単語の読みに基づいてベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体。

【請求項 12】

音声入力装置を備える音声認識装置上で実施される認識文字列の特定処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) 第1乃至第N（2以上のNは自然数）の認識対象文字列によって構成される認識対象文字列群を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 第i（iはN以下の自然数）の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記単語の読みに基づいてベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

(e) 前記第iの認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記音声認識文字列を前記第iの認識対象文字列と比較することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体。

【請求項 13】

前記認識対象文字列群は、アプリケーションから取得されたものであり、前記第 i の認識対象文字列に対応した制御メッセージを生成し、前記アプリケーションに送信することを前記音声認識装置に指示するためのプログラムコードをさらに含む請求項 12 に記載の記憶媒体。

【請求項 14】

前記単語の読みには読みスコアが関連付けられて保存されており、また、前記ベースフォームには発音スコアが関連付けられて保存されており、前記読みスコアと前記発音スコアから導かれる関数値が閾値を越えている場合に前記ベースフォームが前記音声辞書に登録されることを特徴とする請求項 12 に記載の記憶媒体。

【請求項 15】

音声入力装置を備える音声認識装置上で実施される認識文字列の特定処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記単語の読みに基づいてベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ベースフォームのスコアを計算することを前記音声認識装置に指示するためのプログラムコードと、

(e) 前記ベースフォームのスコアが閾値を越えている場合、前記単語に対応付けて音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことに

よって入力された音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得することを前記音声認識装置に指示するためのプログラムコードと、

(h) 前記音声認識文字列を前記第 i の認識対象文字列と比較することを前記音声認識装置に指示するためのプログラムコードと、

(i) 前記音声認識文字列と前記第 i の認識対象文字列が完全に一致した場合には、完全に一致した文字列に関連する処理を実行することを前記音声認識装置に指示するためのプログラムコードと、

(j) 前記音声認識文字列の一部と前記第 i の認識対象文字列の一部が一致した場合には、前記閾値を低い値に更新し、第 2 の閾値とすることを前記音声認識装置に指示するためのプログラムコードと、

(k) 前記第 i の認識対象文字列に含まれる未知語を検出することを前記音声認識装置に指示するためのプログラムコードと、

(l) 検出された未知語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(m) 前記未知語の読みに基づいて第 2 のベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(n) 前記第 2 のベースフォームのスコアを計算することを前記音声認識装置に指示するためのプログラムコードと、

(o) 前記第 2 のベースフォームのスコアが第 2 の閾値を越えている場合、前記未知語に対応付けて前記音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体。

【発明の詳細な説明】

【0001】

【産業上の利用分野】

この発明は、音声認識に関するものであり、より詳しくは、表示画面に表示された情報を音声により特定する方式に関する。

【0002】

【従来の技術】

特開平10-320168号公報のように、従来より表示画面上に表示された情報を音声により特定する方法は存在する。しかし、上記方式においては、アプリケーションの中のメニューやボタンなど、またウェブの中のリンクなどの文字列が音声認識システムの認識可能単語に入っていないなければならないという問題があった。

【0003】

メニューなどの文字列であれば静的にすべての文字列を音声認識の辞書に追加しておくことも考えられるが、ウェブのリンクなどは毎日変更されることも多く、このような静的な手法では対処できない。また、必要以上に辞書に単語を追加しておく、認識率の低下や処理速度の低下という新たな問題が発生することとなる。

【0004】

【発明が解決しようとする課題】

本発明の目的は、音声認識辞書に存在しない単語を含む文字列であっても音声による特定が可能な音声認識システムを提供することにある。

【0005】

本発明の他の一つの目的は、認識精度及び処理速度を所定の水準に維持しつつ、必要となる資源の低減された音声認識システムを提供することにある。

【0006】

本発明の他の一つの目的は、ユーザが直感的に把握できる使い勝手のよい音声認識システムを提供することにある。

【0007】

【課題を解決するための手段】

複数の認識対象文字列によって構成される認識対象文字列群がアプリケーションから取得される。各認識対象文字列は、パーシング・ロジックにより、音声認識単位となる単語に分割される。また、各認識対象文字列に含まれる単語は、音声認識辞書に存在していないが、その読みを特定できる1まとまりの単語である

未知語を含んでいるか否かが確認される。未知語を含んでいる場合には、その読みからその発音を特定するベースフォームを取得し、音声認識辞書に登録する。このベースフォームは、ユーザが、複数の認識対象文字列のうちの1つを音声で特定するために発声した声を認識するために用いられる。

【0008】

本発明の一態様においては、

音声入力装置を備える音声認識装置上で実施される音声情報の登録方法であって、

(a) 第1乃至第N (2以上のNは自然数) の認識対象文字列によって構成される認識対象文字列群を取得する段階と、

(b) 第i (iはN以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する段階と、

(c) 前記単語の読みに基づいてベースフォームを取得する段階と、

(d) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録する段階と、

を含む音声情報の登録方法が提供される。

【0009】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置上で実施される認識文字列の特定方法であって、

(a1) 第1乃至第N (2以上のNは自然数) の認識対象文字列によって構成される認識対象文字列群を取得する段階と、

(a2) 第i (iはN以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する段階と、

(a3) 前記単語の読みに基づいてベースフォームを取得する段階と、

(a4) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録する段階と、

を含む登録段階と、

(b1) 前記第iの認識対象文字列に対応する表示をユーザが読み上げたこと

によって入力された音声情報を取得する段階と、

(b 2) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する段階と、

(b 3) 前記音声認識文字列を前記第 i の認識対象文字列と比較する段階と、
を含む認識段階と、

を含む認識文字列の特定方法が提供される。

【0 0 1 0】

本発明の他の一態様においては、

前記認識対象文字列群は、アプリケーションから取得されたものであり、前記第 i の認識対象文字列に対応した制御メッセージを生成し、前記アプリケーションに送信する段階をさらに含む認識文字列の特定方法が提供される。

【0 0 1 1】

本発明の他の一態様においては、

前記単語の読みには読みスコアが関連付けられて保存されており、また、前記ベースフォームには発音スコアが関連付けられて保存されており、前記読みスコアと前記発音スコアから導かれる関数値が閾値を越えている場合に前記ベースフォームが前記音声辞書に登録されることを特徴とする認識文字列の特定方法が提供される。

【0 0 1 2】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置上で実施される認識文字列の特定方法であって、

(a 1) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する段階と、

(a 2) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する段階と、

(a 3) 前記単語の読みに基づいてベースフォームを取得する段階と、

(a 4) 前記ベースフォームのスコアを計算する段階と、

(a 5) 前記ベースフォームのスコアが閾値を越えている場合、前記単語に対

応付けて音声認識辞書に登録する段階と、

を含む登録段階と、

(b 1) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する段階と、

(b 2) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する段階と、

(b 3) 前記音声認識文字列を前記第 i の認識対象文字列と比較する段階と、

(b 4) 前記音声認識文字列と前記第 i の認識対象文字列が完全に一致した場合には、完全に一致した文字列に関連する処理を実行する段階と、

(b 5) 前記音声認識文字列の一部と前記第 i の認識対象文字列の一部が一致した場合には、前記閾値を低い値に更新し、第 2 の閾値とする段階と、

(b 6) 前記第 i の認識対象文字列に含まれる未知語を検出する段階と、

(b 7) 検出された未知語の読みを取得する段階と、

(b 8) 前記未知語の読みに基づいて第 2 のベースフォームを取得する段階と

(b 9) 前記第 2 のベースフォームのスコアを計算する段階と、

(b 10) 前記第 2 のベースフォームのスコアが第 2 の閾値を越えている場合、前記未知語に対応付けて前記音声認識辞書に登録する段階と、

を含む認識段階と、

を含む認識文字列の特定方法が提供される。

【0 0 1 3】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置であって、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する認識文字列特定部と、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する未知語検出部と、

(c) 前記単語の読みに基づいてベースフォームを取得するベースフォーム生成部と、

(d) 前記ベースフォームが前記単語に対応付けて格納される音声認識辞書とを含む音声認識装置が提供される。

【0014】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置であって、

(a1) 第1乃至第N(2以上のNは自然数)の認識対象文字列によって構成される認識対象文字列群を取得する認識文字列特定部と、

(b) 第i(iはN以下の自然数)の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する未知語検出部と、

(c) 前記単語の読みに基づいてベースフォームを取得するベースフォーム生成部と、

(d) 前記ベースフォームを前記単語に対応付けて格納する音声認識辞書と、

(e) 前記第iの認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する音声入力部と、

(f) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する音声認識エンジンと、

を含み、

(a2) 前記認識特定部は、前記音声認識文字列を前記第iの認識対象文字列と比較する音声認識装置が提供される。

【0015】

本発明の他の一態様においては、

前記認識文字列特定部は、前記認識対象文字列群を、アプリケーションから取得し、前記第iの認識対象文字列に対応した制御メッセージを生成し、前記アプリケーションに送信することを特徴とする音声認識装置が提供される。

【0016】

本発明の他の一態様においては、

前記単語の読みには読みスコアが関連付けられて保存されており、また、前記ベースフォームには発音スコアが関連付けられて保存されており、前記読

みスコアと前記発音スコアから導かれる関数値が閾値を越えている場合に前記ベースフォームが前記音声辞書に登録されることを特徴とする音声認識装置が提供される。

【0 0 1 7】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置であって、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得する認識文字列特定部と、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得する未知語検出部と、

(c) 前記単語の読みに基づいてベースフォームを取得し、前記ベースフォームのスコアを計算するベースフォーム生成部と、

(d) 前記ベースフォームのスコアが閾値を越えている場合、前記単語に対応付けて前記ベースフォームに登録する音声認識辞書と、

(e) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得する音声入力部と、

(f) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得する音声認識エンジンと、

を含み、

(a 2) 前記認識文字列特定部は、

前記音声認識文字列を前記第 i の認識対象文字列と比較し、前記音声認識文字列と前記第 i の認識対象文字列が完全に一致した場合には、完全に一致した文字列に関連する処理を実行し、前記音声認識文字列の一部と前記第 i の認識対象文字列の一部が一致した場合には、前記閾値を低い値に更新し、第 2 の閾値とし、前記未知語検出部に第 i の認識対象文字列に含まれる未知語を検出し、検出された未知語の読みを取得することを指示し、前記未知語の読みに基づいて第 2 のベースフォームを取得し、前記第 2 のベースフォームのスコアを計算することを前記ベースフォーム生成部に指示し、前記第 2 のベースフォームのスコアが第 2 の閾値を越えている場合、前記未知語に対応付けて前記音声認識辞書に登録する音

声認識装置が提供される。

【0018】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置上で実施される音声情報の登録処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) 第1乃至第N(2以上のNは自然数)の認識対象文字列によって構成される認識対象文字列群を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 第i(iはN以下の自然数)の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記単語の読みに基づいてベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【0019】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置上で実施される認識文字列の特定処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) 第1乃至第N(2以上のNは自然数)の認識対象文字列によって構成される認識対象文字列群を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 第i(iはN以下の自然数)の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記単語の読みに基づいてベースフォームを取得することを前記音声認

識装置に指示するためのプログラムコードと、

(d) 前記ベースフォームを前記単語に対応付けて音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

(e) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記音声認識文字列を前記第 i の認識対象文字列と比較することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【 0 0 2 0 】

本発明の他の一態様においては、

前記認識対象文字列群は、アプリケーションから取得されたものであり、前記第 i の認識対象文字列に対応した制御メッセージを生成し、前記アプリケーションに送信することを前記音声認識装置に指示するためのプログラムコードをさらに含む記憶媒体が提供される。

【 0 0 2 1 】

本発明の他の一態様においては、

前記単語の読みには読みスコアが関連付けられて保存されており、また、前記ベースフォームには発音スコアが関連付けられて保存されており、前記読みスコアと前記発音スコアから導かれる関数値が閾値を越えている場合に前記ベースフォームが前記音声辞書に登録されることを特徴とする記憶媒体が提供される。

【 0 0 2 2 】

本発明の他の一態様においては、

音声入力装置を備える音声認識装置上で実施される認識文字列の特定処理プログラムを格納した記録媒体であって、

該プログラムは、

(a) 第 1 乃至第 N (2 以上の N は自然数) の認識対象文字列によって構成される認識対象文字列群を取得することを前記音声認識装置に指示するためのプログラムコードと、

(b) 第 i (i は N 以下の自然数) の認識対象文字列に含まれ、かつ、音声認識辞書には存在しない単語の読みを取得することを前記音声認識装置に指示するためのプログラムコードと、

(c) 前記単語の読みに基づいてベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(d) 前記ベースフォームのスコアを計算することを前記音声認識装置に指示するためのプログラムコードと、

(e) 前記ベースフォームのスコアが閾値を越えている場合、前記単語に対応付けて音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

(f) 前記第 i の認識対象文字列に対応する表示をユーザが読み上げたことによって入力された音声情報を取得することを前記音声認識装置に指示するためのプログラムコードと、

(g) 前記ベースフォームを用いて、前記音声情報を認識し、音声認識文字列を取得することを前記音声認識装置に指示するためのプログラムコードと、

(h) 前記音声認識文字列を前記第 i の認識対象文字列と比較することを前記音声認識装置に指示するためのプログラムコードと、

(i) 前記音声認識文字列と前記第 i の認識対象文字列が完全に一致した場合には、完全に一致した文字列に関連する処理を実行することを前記音声認識装置に指示するためのプログラムコードと、

(j) 前記音声認識文字列の一部と前記第 i の認識対象文字列の一部が一致した場合には、前記閾値を低い値に更新し、第 2 の閾値とすることを前記音声認識装置に指示するためのプログラムコードと、

(k) 前記第 i の認識対象文字列に含まれる未知語を検出することを前記音声認識装置に指示するためのプログラムコードと、

(l) 検出された未知語の読みを取得することを前記音声認識装置に指示する

ためのプログラムコードと、

(m) 前記未知語の読みに基づいて第2のベースフォームを取得することを前記音声認識装置に指示するためのプログラムコードと、

(n) 前記第2のベースフォームのスコアを計算することを前記音声認識装置に指示するためのプログラムコードと、

(o) 前記第2のベースフォームのスコアが第2の閾値を越えている場合、前記未知語に対応付けて前記音声認識辞書に登録することを前記音声認識装置に指示するためのプログラムコードと、

を含む記憶媒体が提供される。

【0023】

【実施例】

A. ハードウェア構成

図1は、本発明の音声認識システム100を実施するためのハードウェア構成の概観図である。音声認識システム100は、中央処理装置(CPU)1とメモリ4とを含んでいる。CPU1とメモリ4は、バス2等を介して、補助記憶装置としてのハードディスク装置13、31とを接続してある。フロッピーディスク装置(またはMO28、CD-ROM26、29等の媒体駆動装置26、28、29、30)20はフロッピーディスクコントローラ(またはIDEコントローラ25、SCSIコントローラ27等の各種コントローラ)19を介してバス2へ接続されている。

【0024】

フロッピーディスク装置(またはMO、CD-ROM等の媒体駆動装置26、28、29、30)20には、フロッピーディスク(またはMO、CD-ROM等の媒体)が挿入され、このフロッピーディスク等やハードディスク装置13、ROM14等の記憶媒体には、オペレーティング・システムと協働してCPU等に命令を与え、本発明を実施するためのコンピュータ・プログラムのコードを記録することができ、メモリ4にロードされることによって実行される。このコンピュータ・プログラムのコードは圧縮し、または、複数に分割して、複数の媒体

にまたがって記録することもできる。

【0025】

音声認識システム100は更に、ユーザ・インターフェース・ハードウェアを備えたシステムとすることができ、ユーザ・インターフェース・ハードウェアとしては、例えば、画面位置情報を入力をするためのポインティング・デバイス（マウス、ジョイスティック、トラックボール等）7、キー入力をサポートするキーボード6や、イメージデータをユーザに提示するためのディスプレイ11、12がある。また、スピーカ23は、オーディオ・コントローラ21から音声信号を、アンプ22を介して受領し、音声として出力する。

【0026】

この音声認識システム100は、シリアルポート15およびモデムまたは、トークンリング等の通信アダプタ18等を介して、他のコンピュータ等と通信を行うことができる。

【0027】

本発明は、通常のパーソナルコンピュータ（PC）やワークステーション、テレビやFAX等の各種家電製品に組み込まれたコンピュータ、車両や航空機等に搭載されるコンピュータ（カーナビゲーション・システム等）又はこれらの組合せによって実施可能である。ただし、これらの構成要素は例示であり、その全ての構成要素が本発明の必須の構成要素となるわけではない。特に、本発明は、音声による文字情報の特定に関するものであるため、パラレルポート16、シリアルポート15等の構成要素は本発明の一態様においては必須のものではない。

【0028】

音声認識システム100のオペレーティング・システムとしては、Windows NT（マイクロソフトの商標）、Windows 9x（マイクロソフトの商標）、Windows 3.x（マイクロソフトの商標）、OS/2（IBMの商標）、Mac OS（Apple社の商標）、Linux（Linus Torvaldsの商標）、AIX（IBMの商標）上のX-WINDOWシステム（MITの商標）などの、標準でGUIマルチウインドウ環境をサポートするものや、PC-DOS（IBMの商標）、MS-DOS（マイクロソフトの商標）などの

キャラクタ・ベース環境のもの、さらにはOS/Open (IBMの商標)、VxWorks (Wind River Systems, Inc.の商標)等のリアルタイムOS、JavaOS等のネットワークコンピュータに組み込まれたOSでも実現可能であり、特定のオペレーティング・システム環境に限定されるものではない。

【0029】

B. システム構成

図2は、本発明の好適な実施例における音声認識システムの処理要素を示す機能ブロック図である。

【0030】

本発明の好適な実施例における音声認識システムは、認識文字特定部201、音声認識エンジン203、未知後検出部205、ベースフォーム生成部207、音声入力部209、アプリケーション211、音声認識辞書231、未知後検出辞書233及び、発音辞書235を備えている。

【0031】

認識文字特定部201は、アプリケーション211から認識対象文字列群を手し、音声認識エンジン203から渡された音声認識文字列に基づいて認識対象文字列群に含まれる認識文字列を特定する。また、未知後検出部205等のコンポーネントを制御する。

【0032】

音声認識エンジン203は、音声認識辞書を用いて、実際に入力された音声情報を解析し、音声認識文字列を出力する。

【0033】

未知語検出部205は、認識文字列特定部201より渡された、認識対象文字列データを受領し、未知語検出辞書233を用いて、未知語を検出し、その読みとスコアを出力する。また、所定のロジックに基づいて、読みスコアの補正を行なう。

【0034】

例えば、表記がかなのみで構成され、確実に長音化されないケースでは、スコアを1に補正する。また、表記にマッチするものが辞書に存在する場合であっても、その辞書の精度が音声認識的には高くない場合（例えばかな漢字変換用の辞書を流用した場合等）、読みが長音化される可能性があればスコアを下げる補正を行なう。この読みスコアは、発生確率等の統計情報、経験値等に基づいて設定することができる。

【0035】

図3は、本発明の好適な実施例における未知語検出辞書233の概念図である。図に示すように、本発明の好適な実施例における未知語検出辞書233には、単語表記301、読み303、発音表記305、読みスコア307の情報が管理されている。

【0036】

ベースフォーム生成部207は、未知語表記及びその読み情報を基に、発音辞書を検索し、対応するベースフォームと発音スコアを出力する。また、所定のロジックに基づいて、発音スコアの補正を行なう。この発音スコアは、発生確率等の統計情報、経験値等に基づいて設定することができる。また、未知語に対するベースフォームのスコアは、読みスコア×発音スコア等、読みスコアと発音スコアに基づいた関数値を設定することができる。

【0037】

図4は、本発明の好適な実施例における発音辞書235の概念図である。図に示すように、本発明の好適な実施例における発音辞書235には、発音表記311、ベースフォーム313、発音スコア315の情報が管理されている。

【0038】

音声入力部209は、ユーザが発声した音声情報をシステムに取り込む。

【0039】

アプリケーション211は、本発明の好適な実施例においては、ウェブ・ブラウザであるが、ワードプロセッサやプレゼンテーション用ソフトウェア等文字情報を扱うことのできるソフトウェアや文字情報に変換可能な画像情報を扱うソフトウェアであればよい。

【0040】

以上図2に示す各機能ブロックを説明したが、これらの機能ブロックは、論理的機能ブロックであり、各々1つのまとまりを持ったハードウェアやソフトウェアによって実現されることを意味するのではなく、複合し、または共通したハードウェアやソフトウェアによって実現可能である。

【0041】

C. 動作手順

本発明の好適な実施例においては、認識文字列の特定のため、大きく分けて以下の4つの動作手順により処理を行なう。

1. 認識対象文字列群の取得
2. 未知語の検出
3. 未知語の発音の音声認識辞書への登録
4. 誤認識時の閾値の動的変動

【0042】

C-1. 認識対象文字列群の取得

この発音スコアは、発生確率等の統計情報、経験値等に基づいて設定することができる。また、未知語に対するベースフォームのスコアは、読みスコア×発音スコア等、読みスコアと発音スコアに基づいた関数値を設定することができる。アプリケーション211がウェブ・ブラウザである場合における認識対象文字列群の取得手順を説明する。

【0043】

まず第1に、MSAA (Microsoft Active Accessibility)を利用する方法 (「MSAA (Microsoft Active Accessibility)」は、マイクロソフト社の商標) が考えられる。ウィンドウズ環境のあるバージョンにおいてはMSAAを利用することが可能である。MSAAで定義されているAPIを使用することによって表示されているブラウザのページのコントロールの情報がリアルタイムに取得できる。このコントロー

ルのなかから、リンクを示すものを抜き出して認識対象文字列群とすることができる。

【0044】

第2に、HTMLのドキュメントを直接読み込む方法が考えられる。この方法においては、ブラウザ上に表示されているページに対応するソースを取得する。そして、そのソースのHTMLタグを解析し、リンクを示すタグの部分の文字列を抽出し、認識対象文字列群とすることができる。

【0045】

第3に、ブラウザの提供するAPIを利用する方法が考えられる。Internet Explorer（「Internet Explorer」は、マイクロソフト社の商標）やNetscape Navigator（「Netscape Navigator」は、ネットスケープ社の商標）などのブラウザは表示されているページの情報を取得するための独自のAPIを提供している。これらのAPIを使用することによって、ページの状態、リンクの情報などを取得できる。

【0046】

以上の方法は、単なる例示であり、本発明の思想に限定を加えるものではない。対象となるアプリケーションから文字列を抽出する手法は種々提案されており、かかる抽出方式の変更は、本発明の実施化に際し当業者が容易におこなうことができるものである。

【0047】

C-2. 未知語の検出

次に、抽出された認識対象文字列から未知語を検出する。ここでいう未知語とは、単語としては認識可能であるが、音声認識辞書231に登録が無い、ベースフォームがシステムにとって未知の単語をいう。

【0048】

図5は、この未知語検出手順を示すフローチャートである。まず、未知語検出部205が認識文字列特定部201から認識対象文字列群（文字列数はN）の最初の認識対象文字列を取得し（ステップ403）、処理を全文字列数繰り返すよ

う設定する（ステップ405）。そして、現在の認識対象文字列を音声認識の単位となる複数の単語に分割する（ステップ407）。

【0049】

この。

英語など単語と単語の間にスペースを入れる言語の場合は、スペースなどの情報を用い単語に分割することは比較的たやすい。しかしながら日本語のようにアジアの漢字圏の言語ではスペースが通常単語と単語の間にはない。このため複雑な単語分割と未知語検索のパーズ（Parse：セグメント化、単語分割）手法を用いる。ナビゲーションに適したパーシング・ロジックについては、形態素解析等、本願の出願時において種々提案されており、当業者に公知の技術であるので、詳細な説明は省略する。

【0050】

そして、パーズされた単語を1つずつチェックし、未知語であると判断された（未知語検出辞書233に存在する）単語を未知語アレイUに登録する（ステップ409～419）。なお、本発明の好適な実施例において、1つの未知語に対し、図3で説明した、読み、発音表記、読みスコアのセットが登録される。

【0051】

以下に、認識対象文字列がパーズされた例を示す。なお「/」は区切り文字であり、下線は未知語である。

例1 「横浜/市営/地下鉄/が/戸塚/から/湘南台/まで/延長」

例2 「IBM/の/e-ビジネス/が/順調/だ」

例3 「Pervasive/ Computing/ Development」

例4 「私/の/名前/は/田原/です」

【0052】

C-3. 未知語に対応するベースフォームの取得・音声認識辞書への登録

次に、未知語に対応するベースフォームの取得及び音声認識辞書への登録の手順を説明する。図6は、この手順を説明するフローチャートである。まず、未知

語が検出された場合（ステップ421）、各未知語の登録処理を行なう（ステップ423～429）。

【0053】

図7、8は、ステップ425の未知語登録サブルーチンの詳細を示すフローチャートである。この手順の概要としては、未知語群に対する発音を予測し、発音をスコア付きで生成する。スコアは発音の確からしさを表す。生成された発音の候補のスコアから認識対象文字列全体のスコアを計算する。計算されたスコアが閾値を超える場合、生成した未知語の発音を動的に認識可能語いとして登録する。なお、ナビゲーション対象文字列のスコアが閾値を超えないような生成発音の場合、登録は行わない。これは精度の低い発音を追加することは認識システム全体の認識率を低下させるおそれがあるためである。

【0054】

この手順をフローチャートに基づき説明する。まず、未知語の数をセットし（ステップ453）、ベースフォーム生成部235が、その未知語の読みから発音辞書235を探索し、未知語に対するベースフォーム群を生成する。各ベースフォームには発音スコアが対応付けられている。なお、発音辞書に対応するベースフォームがない場合には（ステップ457）、エラー処理（ステップ459）を行なう。

【0055】

そして、各認識対象文字列について、発音の組み合わせを求める（ステップ465、467）。図9、10は、認識対象文字列NSに対するn番目の発音の組み合わせを求めるフローチャートである。上記例1の場合、未知語数は2であり、「戸塚」、「湘南台」の読みの組み合わせ、発音の組み合わせがあり、組み合わせは「未知語数」×「読み数」×「発音数」存在する。図10、11においては、この組み合わせを求めている。

【0056】

そして、認識対象文字列全体のスコアを初期設定し（ステップ473）、各未知語のスコアを基に認識対象文字列全体のスコアを計算する（ステップ475～479）。このときの未知語のスコアは、読みスコアおよび発音スコアを基に計

算されている。そして、認識対象文字列全体のスコアが閾値を越えており（ステップ 481）、まだ、その未知語が登録されていない場合には（ステップ 483）、閾値を越えた組み合わせにおけるその未知語のベースフォームを音声認識辞書に登録する（単語表記、ベースフォームの組み合わせで登録する）（ステップ 485）。なお、本発明の好適な実施例においては、未知語の登録は、一時的に行い、認識文字列特定部 201 または音声認識エンジン 203 が停止することにメモリから開放される。

【0057】

この各ステップを、フローチャートに記載された変数を用いて説明すると、まず、認識対象文字列 STR_NAVI の部分文字列としての未知語 $U(i)$ ($1 \leq i \leq n$) に対して生成される発音の候補 $B(i, j(i))$ とする。 $j(i)$ は $U(i)$ に対して生成される発音候補の数 $CN(i)$ に対し、 $1 \leq j(i) \leq CN(i)$ である。発音候補 $B(i, j(i))$ のスコアを $S(i, j(i))$ とする。

【0058】

既知語のスコアを 1 とみなすと、ナビゲーション対象文字列 STR_NAVI 全体のスコア $S(\text{STR_NAVI})$ は、

$$S(\text{STR_NAVI}) = S(1, j(1)) * S(2, j(2)) * \dots * S(n, j(n))$$

となるが、各 $U(i)$ ごとに $CN(i)$ 個のスコアがあるため、 $S(\text{STR_NAVI})$ は

$$SCN = CN(1) * CN(2) * \dots * CN(n)$$

個計算することができる。

【0059】

動的に登録するかを判断する閾値を TH1 とすると、SCN 個の $S(\text{STR_NAVI})$ のうち $S(\text{STR_NAVI}) \geq \text{TH1}$

を満たすような発音群 ($B(1, j(1)), B(2, j(2)), \dots, B(n, j(n))$) を動的に認識辞書に登録する。これにより認識としてはすべてが既知語となるため、ナビゲーション対象文字列 $S(\text{STR_NAVI})$ は認識可能になる。

【0060】

以下に例 1 を用いた例を示す。なお、発音 (Baseform) は「よみ」とは同じではないが、便宜上「よみ」で表現している。

「横浜/市営/地下鉄/が/戸塚/から/湘南台/まで/延長」

「戸塚」

「とつか」 スコア:0.9 S(1,1)

「とづか」 スコア:0.5 S(1,2)

「湘南台」

「しょうなんだい」 スコア:0.9 S(2,1)

【0 0 6 1】

TH1:0.5

$$S(1,1) * S(2,1) = 0.9 * 0.9 = 0.81 \geq 0.5$$

登録(「戸塚(とつか)」 「湘南台(しょうなんだい)」

$$S(1,2) * S(2,1) = 0.5 * 0.9 = 0.45 < 0.5$$

登録しない

【0 0 6 2】

例4の場合の例の場合では、

「私/の/名前/は/田原/です」

「田原」

「たはら」 スコア:0.83 S(1,1)

「たわら」 スコア:0.56 S(1,2)

「たばら」 スコア:0.45 S(1,3)

「たばる」 スコア:0.20 S(1,4)

「だはら」 スコア:0.02 S(1,5)

【0 0 6 3】

TH1:0.5

$$S(1,1) = 0.83 \geq 0.5 \quad \text{登録(「田原(たはら)」)}$$

$$S(1,2) = 0.56 \geq 0.5 \quad \text{登録(「田原(たわら)」)}$$

$S(1,3) = 0.45 < 0.5$ 登録しない

$S(1,4) = 0.20 < 0.5$ 登録しない

$S(1,5) = 0.02 < 0.5$ 登録しない

【 0 0 6 4 】

この詳細なロジックを以下に示す。

RegistBaseform(TH)

begin

STR_NAVIから未知語U(i)を抽出

for i = 1 to N

Baseform生成アルゴリズムによりU(i)から発音候補 B(i,j(i))を生成 ($1 \leq j(i) \leq CN(i)$)

endfor

for i = 1 to N

for k = 0 to CN(i)

j(k) = 1

endfor

endfor

loop = TRUE

while loop == TRUE

S = 1

for i = 0 to N

S = S * S(i,j(i))

endfor

if S >= TH

for i = 0 to N

if B(i,j(i))は発音としてまだ登録されていない

B(i,j(i))を登録

endif

```

endfor
endif
if next_path( 1 ) == TRUE
loop = FALSE
endif
endwhile
end

next_path( i )
begin
if i > N
return TRUE
endif
if next_path( i+1 ) == TRUE
j(i) = j(i) + 1
endif
if j(i) > CN(i)
j(i) = 1
return TRUE
endif
return FALSE
end

```

【 0 0 6 5 】

C - 4 . 誤認識時の閾値の動的変動

この項で説明するご認識時の閾値の動的変動手順は、本発明の追加の手順であり、本発明の一態様においては構成要素とはならない。この手順を図 1 1 のフローチャートを基に説明する。

【0066】

まず、ユーザの発声により入力された音声情報（音声コマンド）を取り込み、音声認識エンジン203を用いて、音声認識文字列を取得する（ステップ553、555）。この例では、説明の便宜上、音声認識エンジンが出力する音声認識文字列は1つのみであるが、音声認識エンジンからも、音声認識スコア付きの複数の音声認識文字列から構成される音声認識群を返してもよい。かかる場合は、この手順が音声認識文字列数繰り返され、ステップ563においては一致度のスコア計算がされる。

【0067】

そして、認識対象文字列数手順を繰り返すために、変数*i*が初期設定される（ステップ561）。そして、比較する音声認識文字列が認識対象文字列の*i*番目に一致するか否かが検査される（ステップ563）。

【0068】

比較する音声認識文字列が認識対象文字列の*i*番目に一致する場合には、その認識対象文字列*i*が音声認識文字列に対応する文字列として認識され、必要な処理のために使用される（ステップ565）。例えば、その認識対象文字列がウェブ・ブラウザから取得された文字列である場合には、その文字列に対応付けられたURLを表示すべきURLとしてウェブ・ブラウザ側に送信したり、ワードプロセッサ等においては、その文字列を反転させたり、その文字列に対応したコマンドを実行したりすることができる。

【0069】

この一方、比較する音声認識文字列が認識対象文字列の*i*番目に一致しない場合には、先頭から未知語部分までの一致があるか判断される（ステップ567）。一致がない場合には、この認識対象文字列は、音声認識文字列に対応しない文字列であると判断し、次の認識対象文字列の検査を行なう（ステップ573）。なお、もう認識対象文字列が存在しない場合には、認識エラーを表示し（ステップ572）、再度音声コマンドを入力するようにユーザに指示する。

【0070】

先頭から未知語部分までの一致があると判断された場合には、未知語のベース

フォームの登録時に比較される閾値を下げ（ステップ 5 6 9）、その認識対象文字列に含まれる未知語の登録を行なう（ステップ 5 7 1）。登録された未知語は、次の音声認識に使用される。この手順により、認識エラーが未知語の発音により引き起こされている可能性がある場合、閾値の動的変動を行うことで、認識エラーが多発することを防止している。

【 0 0 7 1 】

例えば、「私/の/名前/は/田原/です」の例の場合TH1が0.5であるため、動的に登録された発音は「たはら」と「たわら」である。ところが話者は「田原」に対し「たばら」と発音する場合、認識エンジンはこの文字列全体を認識できず、「私/の/名前/は/」以降のパスは返されない。こういった場合、先頭からのパスが一致ものの閾値を見直し（下げる）、再度、動的に発音を追加することによって認識対象とすることができる。

【 0 0 7 2 】

前記例 4 を用いて説明を行なう。この手順においては、TH1:0.5では「私/の/名前/は」までしか認識しないかった場合、TH2:0.25として再度登録する。

「私/の/名前/は/田原/です」

【 0 0 7 3 】

「田原」

「たはら」 スコア:0.83 S(1,1)

「たわら」 スコア:0.56 S(1,2)

「たばら」 スコア:0.45 S(1,3)

「たばる」 スコア:0.20 S(1,4)

「だはら」 スコア:0.02 S(1,5)

【 0 0 7 4 】

TH2:0.25

S(1,1) = 0.83 >= 0.25 登録済み

S(1,2) = 0.56 >= 0.25 登録済み

S(1,3) = 0.45 >= 0.25 登録（「田原（たばら）」）

S(1,4) = 0.20 < 0.25 登録しない

S(1,5) = 0.02 < 0.25 登録しない

【 0 0 7 5 】

この誤認識時の閾値の動的変動手順のロジックを以下に示す。

for i = 1 to N

NAVI_STR(i)とリジェクトされた認識結果を比較

if 先頭から未知語部分まで一致する

TH = TH2 /* 前より低い閾値 */

RegistBaseform(TH) /* 再度発音を登録 */

endif

endfor

【 0 0 7 6 】

【発明の効果】

以上説明したように、本発明によれば、音声認識辞書に存在しない単語を含む文字列であっても音声による特定が可能となる。

【 0 0 7 7 】

【図面の簡単な説明】

【図 1】 本発明の音声認識装置のハードウェア構成の一実施例を示すブロック図である。

【図 2】 本発明の好適な実施例における音声認識システムの処理要素のブロック図である。

【図 3】 本発明の好適な実施例における未知語検出辞書の概念図である。

【図 4】 本発明の好適な実施例における発音辞書の概念図である。

【図 5】 本発明の好適な実施例における未知語検出手順を示すフローチャートである。

【図 6】 本発明の好適な実施例における未知語に対応するベースフォームの取得及び音声認識辞書への登録の手順を示すフローチャートである。

【図 7】 本発明の好適な実施例における未知語登録サブルーチンの詳細を示すフローチャートである。

【図 8】 本発明の好適な実施例における未知語登録サブルーチンの詳細を示すフローチャートである。

【図 9】 本発明の好適な実施例における認識対象文字列に対する発音の組み合わせを求めるフローチャートである。

【図 1 0】 本発明の好適な実施例における認識対象文字列に対する発音の組み合わせを求めるフローチャートである。

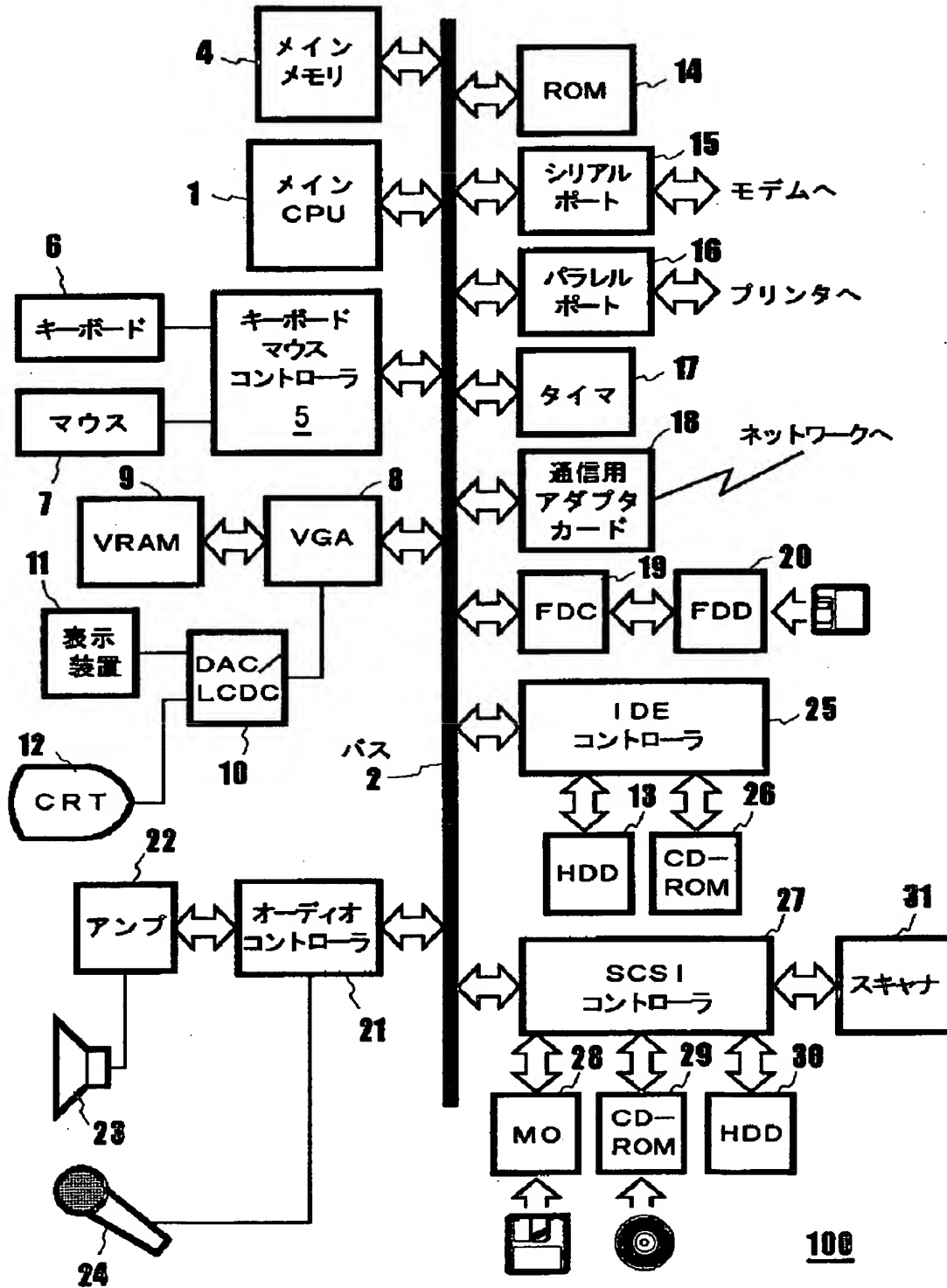
【図 1 1】 本発明の好適な実施例における音声認識手順を示すフローチャートである。

【符号の説明】

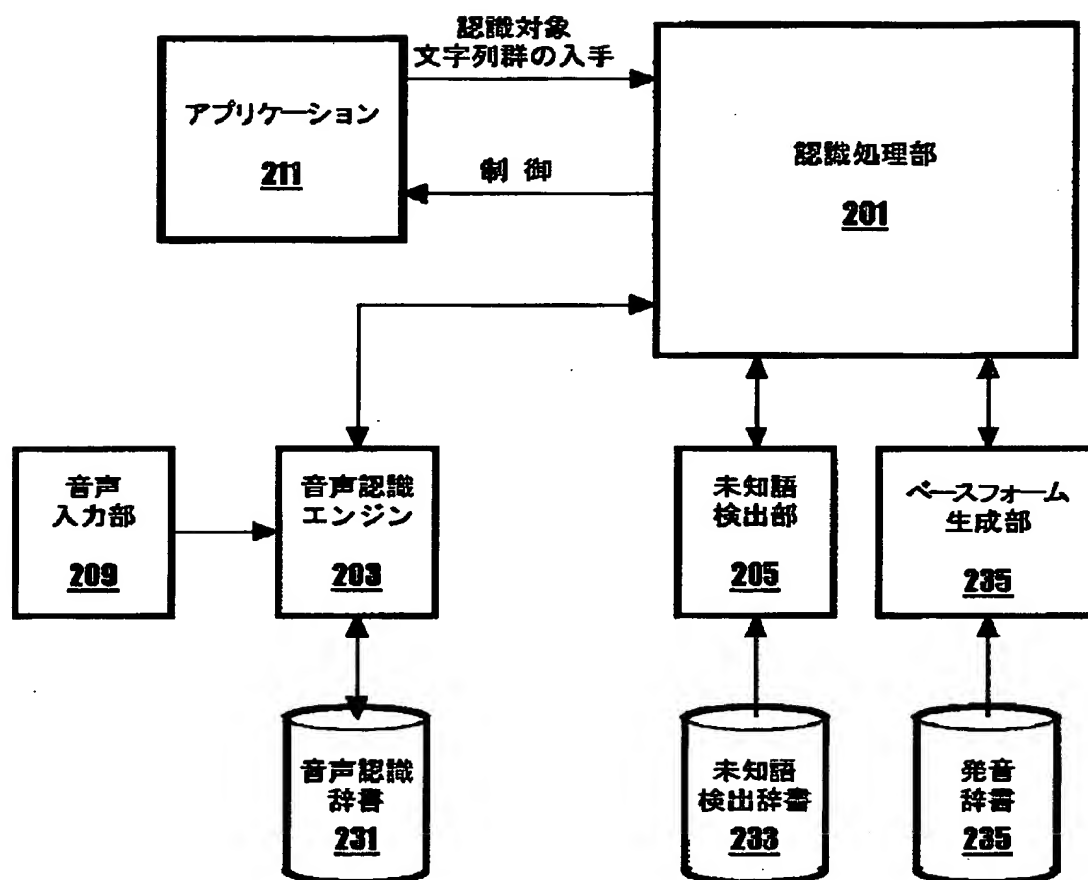
- 1 0 0 音声認識システム
- 2 0 1 認識文字列特定部
- 2 0 3 音声認識エンジン
- 2 0 5 未知語検出部
- 2 0 7 ベースフォーム生成部
- 2 0 9 音声入力部
- 2 1 1 アプリケーション
- 2 3 1 音声認識辞書
- 2 3 3 未知語検出辞書
- 2 3 5 発音辞書

【書類名】 図面

【図 1】



【図 2】



【図 3】

単語表記 <u>301</u>		
田 原		
読み <u>303</u>	発音表記 <u>305</u>	読みスコア <u>307</u>
たはら	tahara	0.83
たわら	tawara	0.56
たばら	tabara	0.45
たばる	tabaru	0.20
だはら	dahara	0.02

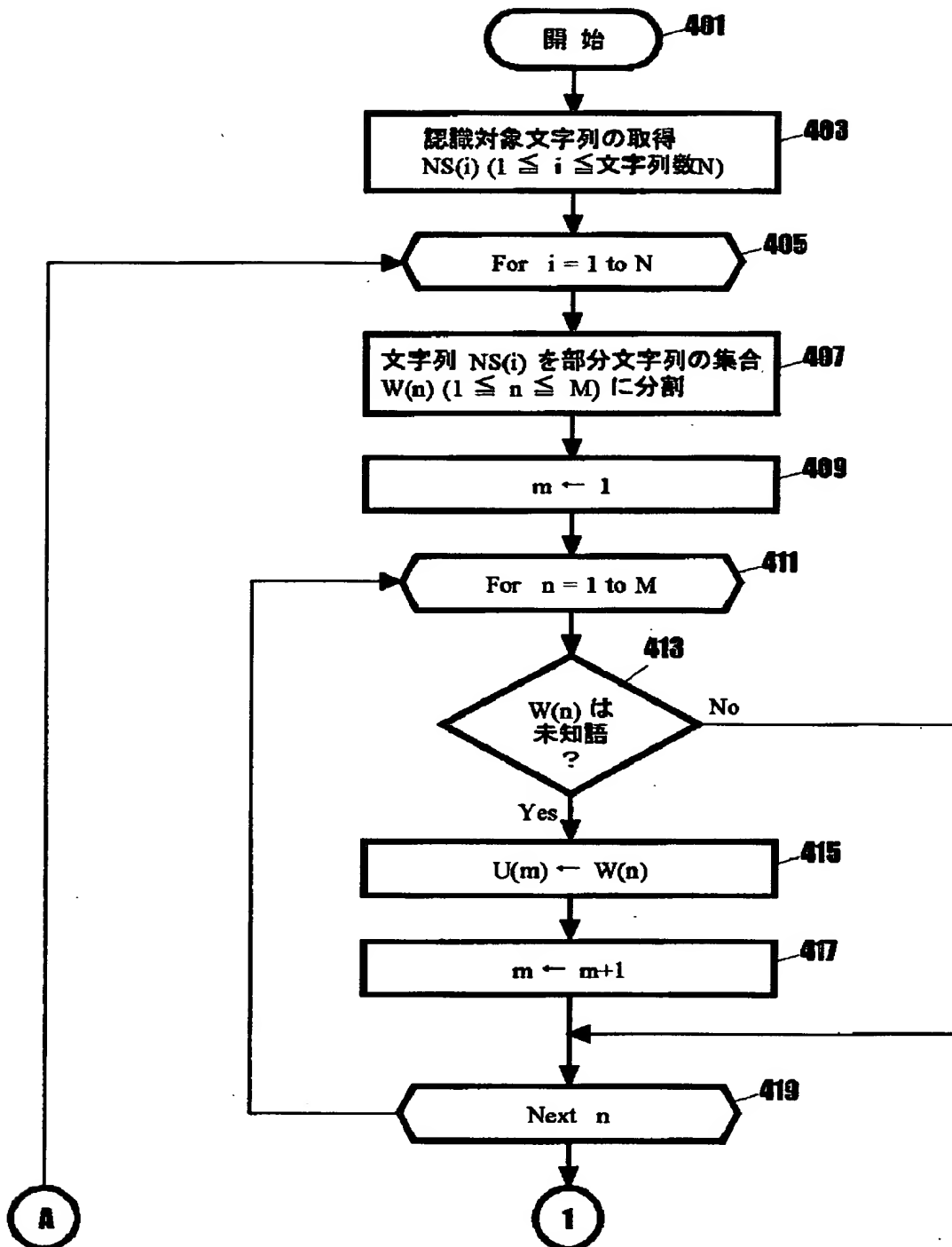
300

【図 4】

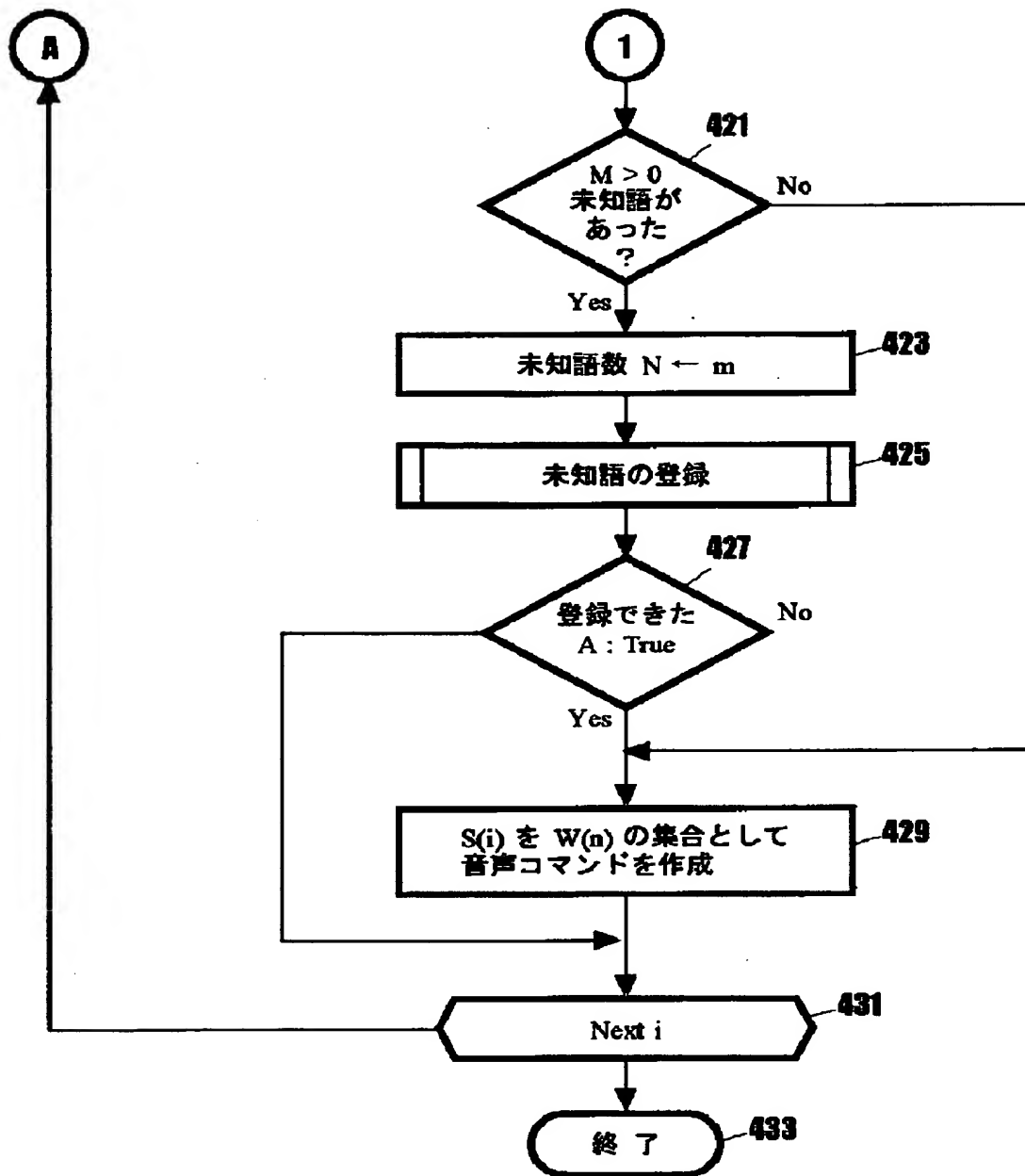
発音表記 <u>311</u>	
こ う し	
ベースフォーム <u>313</u>	発音スコア <u>315</u>
ベースフォーム 1	0.74
ベースフォーム 2	0.58
ベースフォーム 3	0.37

310

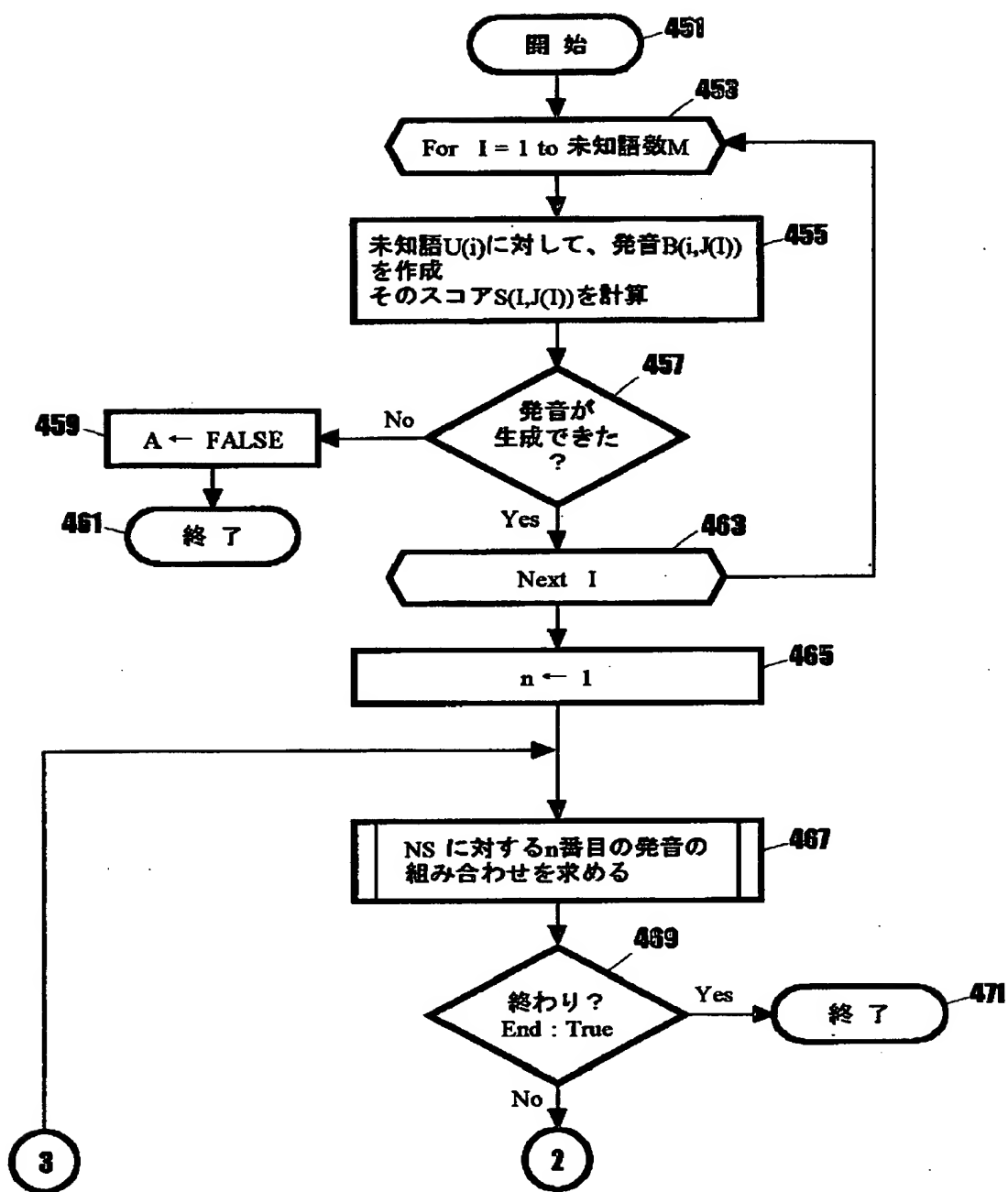
【図 5】



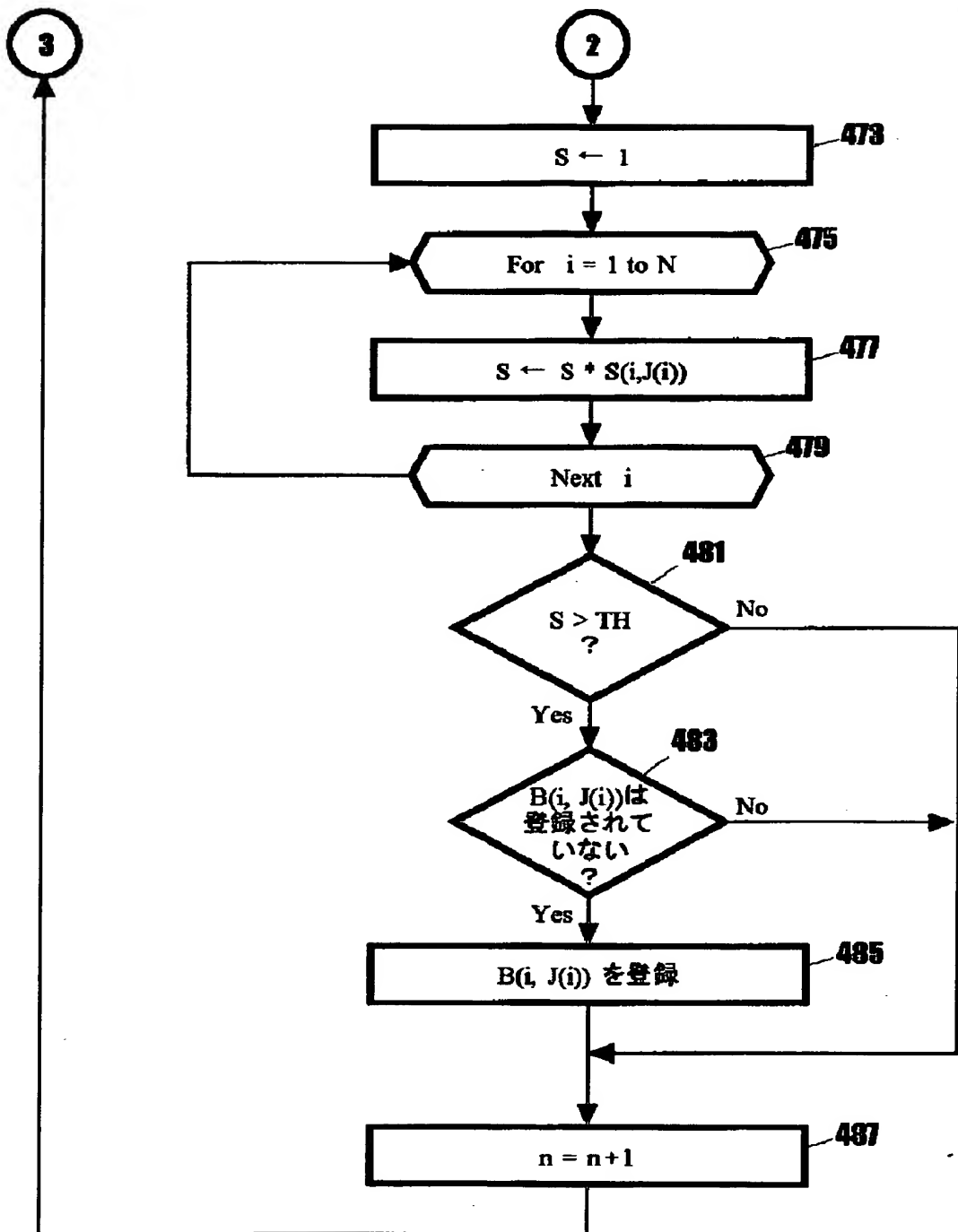
【図 6】



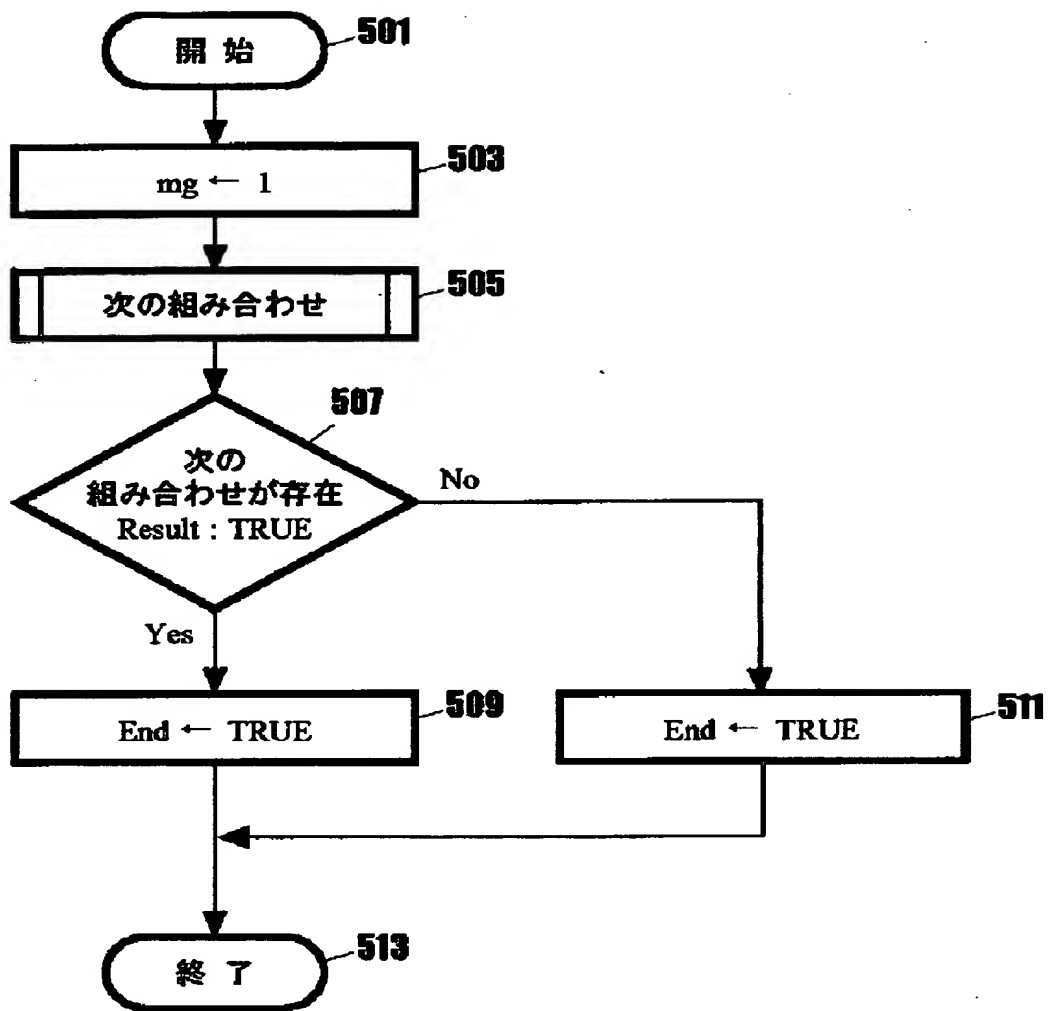
【図 7】



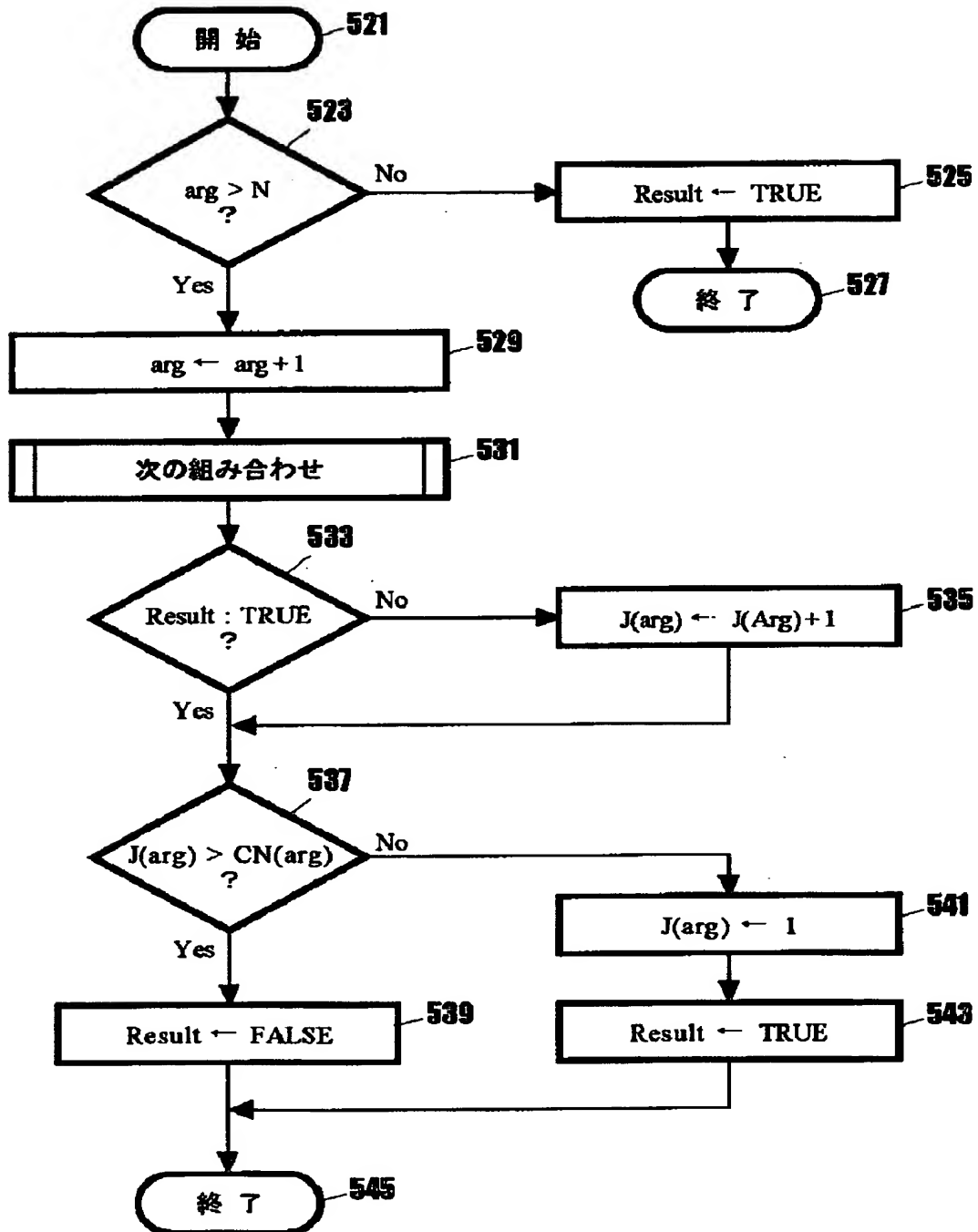
【図 8】



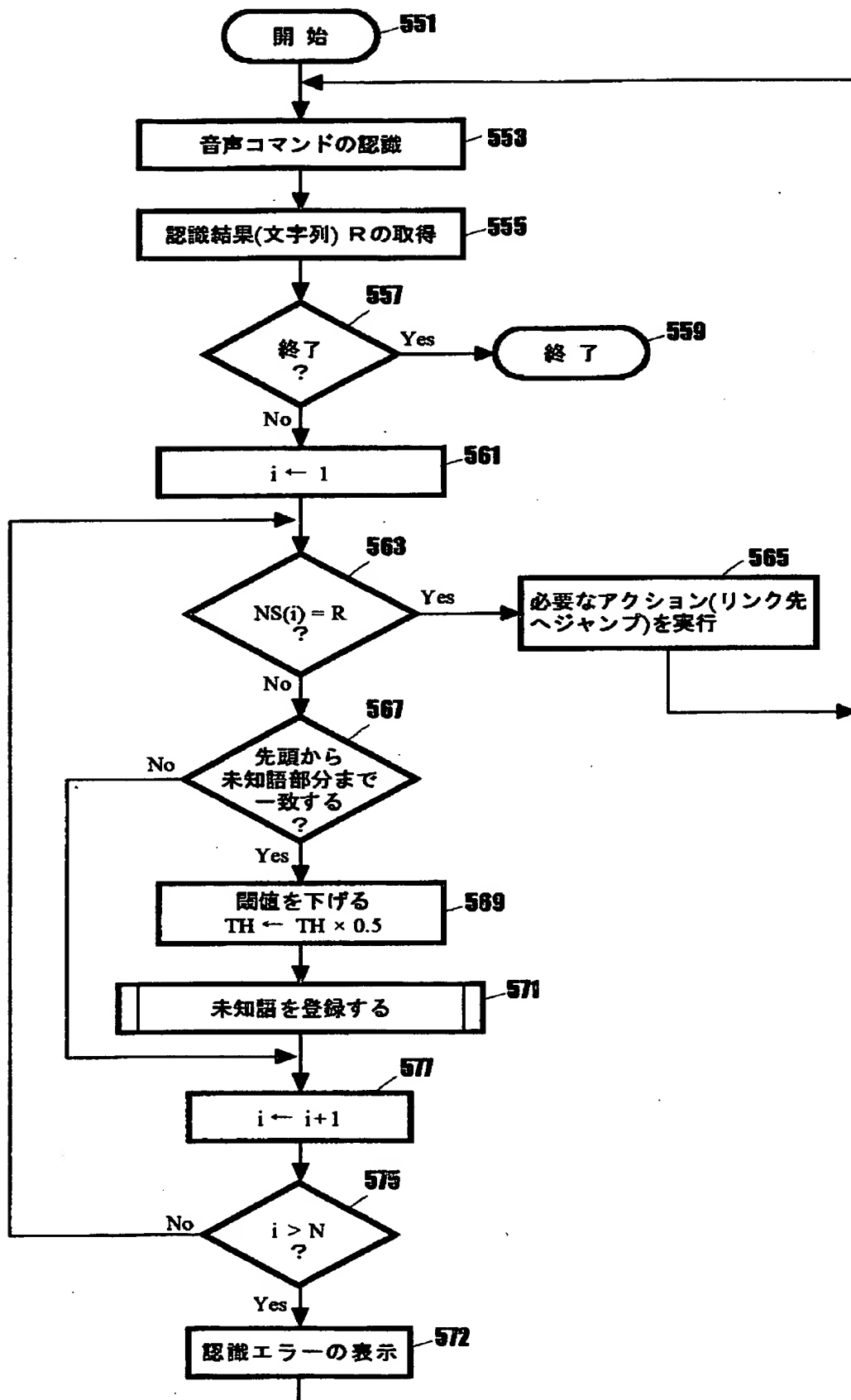
【図 9】



【図 10】



【図 11】



【書類名】 要約書

【要約】

【課題】 音声認識辞書に存在しない単語を含む文字列であっても音声による特定を可能にする。

【解決手段】

複数の認識対象文字列によって構成される認識対象文字列群がアプリケーションから取得される。各認識対象文字列は、パーシング・ロジックにより、音声認識単位の単語に分割される。また、各認識対象文字列に含まれる単語は、音声認識辞書に存在していないが、その読みを特定できる1まとまりの単語である未知語を含んでいるか否かが確認される。未知語を含んでいる場合には、その読みからその発音を特定するベースフォームを取得し、音声認識辞書に登録する。このベースフォームは、ユーザが、複数の認識対象文字列のうちの1つを音声で特定するために発声した声を認識するために用いられる。

【選択図】 図2

認定・付加情報

特許出願の番号	平成11年 特許願 第257587号
受付番号	59900885512
書類名	特許願
担当官	第七担当上席 0096
作成日	平成11年 9月14日

<認定情報・付加情報>

【提出日】	平成11年 9月10日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [390009531]

1. 変更年月日 1990年10月24日

[変更理由] 新規登録

住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)

氏 名 インターナショナル・ビジネス・マシーンズ・コーポレイション